SHIELD PROTOCOL 10.1

THE LAMBDA STANDARD

Statistical Hierarchy for Identifying Evidence of
Literary Design in Sacred Texts

Version 10.1 Specification

January 30, 2026

Royal Levitical Priesthood Ministry
Hearts2Fathers Publishing

Principal Researcher: Dallas James

TABLE OF CONTENTS

1. Executive Summary

SHIELD Protocol 10.1 (Lambda Standard) represents a methodologically refined iteration of the Kappa Standard (10.0), incorporating precision corrections for statistical edge cases while preserving the core analytical framework. This version addresses mathematical boundary conditions identified during rigorous peer review, implementing exact combinatorial methods for small-sample scenarios and density-normalized baselines for high-saturation passages.

1.1 Upgrade Philosophy

The Lambda Standard maintains full backward compatibility with Kappa verdicts while providing enhanced precision in three critical domains:

Continuity Correction: CSI calculations now incorporate the Yates continuity correction for improved accuracy with discrete data, reducing Type I errors in marginal cases.

Exact Critical Values: Small-sample passages (n_1 or $n_2 < 10$) now utilize Wald-Wolfowitz lookup tables rather than Z-score approximations, eliminating normal-distribution assumption violations.

Density-Normalized RLM: Run Length Metric baselines now account for marker density through Monte Carlo simulation, preventing false Dominant ratings in naturally saturated passages.

1.2 Retained Framework

The Dominion Trinity theological framework (Identity, Inheritance, Priesthood) remains the foundational marker system. The SSI weighting for these markers is intentional and constitutive of the Protocol's purpose—detecting divine architectural patterns—not a bias requiring correction.

2. Mathematical Corrections (Lambda Patches)

2.1 CSI-C: Continuity-Corrected Clustering Strength Index

The original CSI formula derived a Z-score from the Wald-Wolfowitz Runs Test:

$$CSI = (R_obs - E(R)) / \sigma$$

For discrete data (verses are countable units), the continuous normal approximation introduces systematic error. The Lambda Standard implements continuity correction:

$$CSI\text{-}C = (|R_obs - E(R)| - 0.5) / \sigma$$

Where:

R_obs = Observed number of runs (alternations between marked/unmarked sequences)

E(R) = Expected runs = $(2 \times n_1 \times n_2) / N + 1$

σ = Standard deviation = $\sqrt{[(2 \times n_1 \times n_2 \times (2 \times n_1 \times n_2 - N)) / (N^2 \times (N - 1))]}$

n_1 = Marked verses

n_2 = Unmarked verses

N = Total verses ($n_1 + n_2$)

The subtraction of 0.5 from the absolute difference compensates for the discrete-to-continuous approximation, reducing false positive rates by approximately 3-5% in passages with 20-50 verses.

2.2 Exact Critical Value Protocol

When $\min(n_1, n_2) < 10$, the CSI-C formula becomes unreliable due to insufficient sample size for normal approximation. In these cases, Protocol 10.1 mandates consultation of exact Wald-Wolfowitz critical value tables.

2.2.1 Implementation Procedure

1. Calculate n_1 (marked) and n_2 (unmarked) for the passage.

2. If $\min(n_1, n_2) \geq 10$: Use CSI-C formula.

3. If $\min(n_1, n_2) < 10$: Consult Appendix A lookup table.

4. Compare R_obs against critical values for $\alpha = 0.05$ (two-tailed).

5. If R_obs ≤ Lower Critical or R_obs ≥ Upper Critical: Reject null hypothesis (significant clustering or dispersion).

2.3 RLM-D: Density-Normalized Run Length Metric

The Kappa Standard's RLM used the Schilling approximation for expected longest run:

$$E_max \approx log(N) / log(1/p)$$

This formula assumes moderate probability (p ≈ 0.5) and fails catastrophically at high densities (p > 0.7). For a passage with 81.5% marker density, the formula predicts E_max ≈ 4.2 when the actual expected longest run approaches 16 verses.

2.3.1 Monte Carlo Baseline Method

Protocol 10.1 replaces the Schilling approximation with empirically-derived baselines:

1. For a passage with N verses and M markers, generate 10,000 random permutations.
2. For each permutation, calculate the longest consecutive run of marked verses.
3. The mean of these 10,000 values = E_max (empirical).
4. The standard deviation of these values = σ_max.

$$RLM\text{-}D = (L_obs - E_max) / \sigma_max$$

Where L_obs = observed longest run.

2.3.2 Pre-Computed Lookup Tables

To avoid computational overhead, Appendix B provides pre-computed E_max and σ_max values for common (N, M) combinations. For unlisted combinations, the Monte Carlo method should be executed.

2.3.3 Density Tier Adjustment

The Lambda Standard introduces Density Tiers to contextualize RLM-D verdicts:

Density (p)	Tier	RLM-D Interpretation
< 30%	Sparse	Standard interpretation applies
30-60%	Moderate	Standard interpretation applies
60-80%	Dense	Requires NRI context for Dominant verdict
> 80%	Saturated	RLM-D primarily measures efficiency; NRI carries structural weight

At Saturated density, a passage may receive an Expected RLM-D yet still qualify as Monumental if the NRI itself is statistically extraordinary. The markers are inevitable, but the choice to place them in that passage is not.

3. Enhanced Metrics

3.1 CMC-ϕ: Phi Coefficient Co-occurrence

The Lambda Standard retains the Ochiai Coefficient (CMC) as the primary co-occurrence measure while adding the Phi Coefficient as a chance-corrected secondary metric:

$$CMC\ (Ochiai) = |A \cap B| \ / \ \sqrt{(|A| \times |B|)}$$

$$CMC\text{-}\phi\ (Phi) = (n_{11} \times n_{00} - n_{10} \times n_{01}) \ / \ \sqrt{(n_{1\bullet} \times n_{0\bullet} \times n_{\bullet 1} \times n_{\bullet 0})}$$

Where the contingency table is:

	Marker B Present	Marker B Absent
Marker A Present	n_{11}	n_{10}
Marker A Absent	n_{01}	n_{00}

CMC-ϕ ranges from -1 (perfect negative association) through 0 (chance-level co-occurrence) to +1 (perfect positive association). Unlike Ochiai, Phi subtracts expected random overlap from observed overlap.

3.1.1 Reporting Protocol

When passage density exceeds 50% for either marker:

• Report both CMC and CMC-ϕ

• Flag if CMC > 0.7 but CMC-ϕ < 0.3 (density-inflated co-occurrence)

• Use CMC-ϕ for theological pairing conclusions

3.2 Chiasmus Detection: Dual Algorithm Approach

Protocol 10.1 retains Levenshtein Distance as the primary chiasmus detector while adding Smith-Waterman local alignment for embedded pattern discovery.

3.2.1 Primary: Levenshtein-Based Chiasmus Score

$$ChiScore = 1 - (Lev(S, S_rev) \ / \ max(|S|, |S_rev|))$$

Appropriate for: Whole-pericope analysis where the entire passage is expected to form the chiastic structure.

3.2.2 Secondary: Smith-Waterman Embedded Detection

For passages exceeding 20 verses, the Lambda Standard adds an optional Smith-Waterman scan:

• Align marker sequence S against its reverse S_rev

• Scoring: Match = +2, Mismatch = -1, Gap = -1

• Minimum alignment length: 5 elements

• Report local alignment score and position range

This method detects partial chiasms embedded within larger non-chiastic contexts (e.g., a 7-verse chiasm within a 30-verse chapter).

3.2.3 Chiasmus Verdict Hierarchy

Verdict	Criteria	Interpretation
Perfect Chiasm	ChiScore ≥ 0.95	Complete inverted parallelism
Strong Chiasm	ChiScore 0.80-0.94	Clear pattern with minor variation
Moderate Chiasm	ChiScore 0.60-0.79	Discernible pattern with gaps
Embedded Chiasm	SW Score ≥ 10, ChiScore < 0.60	Local chiasm within larger passage
No Chiasm	ChiScore < 0.60, SW < 10	No detectable chiastic structure

4. Retained Metrics (Unchanged from Kappa)

4.1 NRI: Non-Random Index

The foundational saturation metric remains unchanged:

$$NRI = (O - E) / \sigma_E$$

Where O = observed marker count, E = expected count under random distribution, σ_E = standard deviation of expectation.

4.2 SDS: Standardized Deviance Score

Passage-level deviance measurement:

$$SDS = (P_{obs} - P_{exp}) / SE$$

Where P_obs = observed marker proportion, P_exp = corpus-wide proportion, SE = standard error.

4.3 Inclusio Detection

$$InclusioScore = 1 - (|M_{first} - M_{last}| / \max(M_{first}, M_{last}))$$

Detects bookending structures where similar markers frame passage boundaries.

4.4 Climactic Index

Centroid-based method for detecting front-loading versus back-loading of markers:

$$CI = (Centroid_{obs} - Centroid_{uniform}) / \sigma_{centroid}$$

Positive values indicate back-loaded (climactic) arrangement; negative values indicate front-loaded structure.

4.5 SSI: Structural Signature Index

Composite metric incorporating Dominion Trinity co-occurrence:

$$SSI = 0.4 \times CSI + 0.3 \times RLM + 0.2 \times CMC_{DT} + 0.1 \times (Chi + Inc)$$

Where CMC_DT = average co-occurrence of Identity, Inheritance, and Priesthood markers.

Note: The Dominion Trinity weighting is constitutive to the Protocol's design purpose. Parameterization would convert SHIELD from a discovery instrument into a confirmation tool.

5. Verdict Framework

5.1 Structural Classification

Verdict	Criteria	Interpretation
Random	SSI < 1.5	No detectable non-random structure
Patterned	SSI 1.5-2.5	Weak structural signals present
Architectural	SSI 2.5-3.5	Clear intentional design evident
Monumental	SSI > 3.5	Extraordinary structural integration

5.2 Lambda-Specific Qualifiers

Protocol 10.1 introduces mandatory qualifiers for edge cases:

[Small-Sample]: Applied when $\min(n_1, n_2) < 10$. Indicates exact-value method was used.

[Density-Adjusted]: Applied when marker density > 60%. Indicates RLM-D baseline was Monte Carlo derived.

[Embedded-Chi]: Applied when Smith-Waterman detected local chiasm not visible in global Levenshtein score.

[Phi-Corrected]: Applied when CMC-ϕ diverges significantly from CMC, indicating density-inflated co-occurrence.

6. Worked Example: Revelation 21 (Lambda Recalculation)

6.1 Passage Parameters

New Jerusalem Vision | Revelation 21:1-27

Total Verses (N): 27

Inheritance Markers (n_1): 22

Unmarked Verses (n_2): 5

Density (p): 81.5%

Observed Runs (R_obs): 4

Observed Longest Run (L_obs): 15

6.2 CSI-C Calculation

Since $n_2 = 5 < 10$, we must use exact critical values (Section 2.2).

From Appendix A, for $n_1 = 22$, $n_2 = 5$:

Lower Critical Value: 4

Upper Critical Value: 12

R_obs = 4 equals the lower critical value.

Verdict: Significant clustering at α = 0.05 (borderline). Qualifier: [Small-Sample]

6.3 RLM-D Calculation

With density = 81.5%, we apply Monte Carlo baseline (Appendix B):

E_max (simulated): 16.1 verses

σ_max (simulated): 2.8 verses

$$RLM\text{-}D = (15 - 16.1) / 2.8 = -0.39$$

The observed longest run (15) is slightly below the expected value for this density level.

Verdict: Expected/Random structure given saturation level. Qualifier: [Density-Adjusted]

6.4 Integrated Interpretation

The Lambda Standard reveals a nuanced picture of Revelation 21:

• NRI: Extraordinary saturation (22/27 = 81.5% vs corpus baseline ~15%)

• CSI-C: Borderline significant clustering [Small-Sample]

• RLM-D: Expected run length for density level [Density-Adjusted]

Compound Verdict: The passage exhibits Monumental thematic saturation (NRI-driven) with density-appropriate spatial arrangement. The structural signature emerges from the choice to concentrate Inheritance markers at this locus, not from unusual clustering within that concentration.

7. Version History and Compatibility

Version	Codename	Key Features
7.0	Foundation	NRI, FDR correction, empirical baselines, genre sensitivity
8.0	Alpha	SDS integration, effect-size tiers
9.0	Beta	Multi-corpus validation, passage-level analysis
10.0	Kappa	Spatial metrics (CSI, RLM, CMC), chiasmus detection, SSI
10.1	Lambda	Continuity correction, exact critical values, density-normalized RLM, Phi coefficient, Smith-Waterman option

Backward Compatibility: All Kappa (10.0) verdicts remain valid. Lambda refinements affect only edge cases (small samples, high density). Re-analysis is recommended for passages with $n_2 < 10$ or density > 70%.

Appendix A: Wald-Wolfowitz Exact Critical Values

For use when $\min(n_1, n_2) < 10$. Reject H_0 (random arrangement) if $R_{obs} \leq$ Lower or $R_{obs} \geq$ Upper at $\alpha = 0.05$.

n_1	n_2	Lower Critical	Upper Critical	N
5	5	2	10	10
10	5	3	10	15
15	5	3	11	20
20	5	4	11	25
22	5	4	12	27
25	5	4	12	30
10	10	6	16	20
15	10	7	18	25
20	10	9	18	30

Extended tables available in supplementary materials. For unlisted combinations, consult Swed & Eisenhart (1943) or compute exact probabilities.

Appendix B: Monte Carlo RLM Baselines

Pre-computed expected longest run (E_max) and standard deviation (σ_max) for common (N, M) combinations. Based on 10,000 random permutations per configuration.

N	M (marked)	Density	E_max	σ_max	Tier
27	8	29.6%	2.9	1.2	Sparse
27	14	51.9%	5.1	1.9	Moderate
27	18	66.7%	8.4	2.4	Dense
27	22	81.5%	16.1	2.8	Saturated
50	15	30.0%	3.4	1.3	Sparse
50	25	50.0%	5.8	1.8	Moderate
50	35	70.0%	10.2	2.6	Dense
50	42	84.0%	22.7	4.1	Saturated
100	30	30.0%	4.1	1.4	Sparse
100	50	50.0%	6.9	1.9	Moderate
100	70	70.0%	12.8	2.9	Dense
100	85	85.0%	38.4	6.2	Saturated

For unlisted (N, M) combinations, execute Monte Carlo simulation with 10,000 permutations.

Appendix C: Implementation Checklist

Analysts should verify the following before issuing Lambda Standard verdicts:

☐ Passage boundaries confirmed (pericope integrity)

☐ Marker tagging verified against UTM_LFE_corpus_7_0.jsonl

☐ Sample size check: If $\min(n_1, n_2) < 10$, use exact critical values

☐ Density tier identified

☐ If density > 60%, Monte Carlo baseline applied for RLM

☐ If density > 50% for either co-occurrence marker, report both CMC and CMC-ϕ

☐ If passage > 20 verses, Smith-Waterman scan completed

☐ All applicable qualifiers appended to verdict

☐ Compound interpretation addresses NRI-structure relationship

— End of Specification —